Journal of Benefit-Cost Analysis (2024), 15: S1, 26–51
doi:10.1017/bca.2024.6

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Social Welfare Functions and Health Policy: A New Approach

Matthew D. Adler

Duke University, Durham, NC, USA
Email: adler@law.duke.edu

Keywords: prioritarianism; social value of risk reduction; social welfare function; utilitarianism; value of statistical life

Abstract

The social welfare function (SWF) framework converts the possible outcomes of governmental policy choice into vectors (lists) of interpersonally comparable well-being numbers, measuring the lifetime well-being of each individual in the population of interest. The SWF proper is a rule for ranking these vectors. The utilitarian SWF adds up well-being numbers. A prioritarian SWF adds up well-being numbers plugged into a strictly increasing and strictly concave transformation function. Governmental policies are conceptualized as probability distributions over well-being vectors. A recent literature applies the SWF framework to health policy. This article first provides a brief overview of the SWF framework and then reviews some of the key concepts and findings that have emerged from this literature. One such concept is the "social value of risk reduction" (SVRR): the marginal social value (as calculated by the SWF) per unit of reduction in fatality risk for a given individual. The SVRR is the analogue, within the SWF framework, to the value-of-statistical-life (VSL) concept within benefit–cost analysis. This article explicates the SVRR concept and reports on recent theoretical findings and simulations that illustrate the properties of utilitarian and prioritarian SVRRs and their differences from VSL.

1. Introduction

The social welfare function (SWF) framework originates in theoretical welfare economics (Bossert & Weymark, 2004; Weymark, 2016) and is now widely used in some policy-focused literatures, such as optimal taxation (Tuomala, 2016) and climate economics (Botzen & van den Bergh, 2014). It has also been employed to assess health policy. The SWF literature in health economics can be divided into two branches. An earlier branch applies the SWF to some measure of individual health, such as a QALY (Bleichrodt *et al.,* 2004; Dolan, 1998; Hougaard *et al.,* 2013; Østerdal, 2005; Williams, 1997). A more recent branch evaluates health policy by applying the SWF to a measure of individual *well-being* (Adler, 2017, 2019, chap. 5, 2020a, 2020b, Forthcoming; Adler *et al.,* 2014, 2021, 2023; Cookson *et al.,* 2022; Ferranna *et al.,* 2022, 2023; Hammitt & Treich, 2022). The inputs to the well-being measure are *all* the individual attributes that determine individual welfare (up to the limits of modeling tractability), including not merely health but also the individual's material resources (income, consumption, or wealth).

Section 2 of this article provides a brief overview of the SWF methodology in general. Section 3 reviews some key concepts and findings that have emerged in this second, newer, branch of SWF-based health economics.

The reader may well wonder why the SWF framework should be used in lieu of benefit–cost analysis (BCA). I have elsewhere addressed this question at length (Adler, 2012, 2017, 2019); space constraints preclude recapitulating my arguments here. Section 3 will note differences between BCA and the SWF methodology with respect to health policy, but a normative defense of the SWF approach will not be undertaken in this article.

By "BCA," I mean the traditional version: ranking policies according to the unweighted sum of individuals' monetary equivalents (compensating or equivalent variations). A different version sums monetary equivalents multiplied by distributional weights (Adler, 2016a; Boadway, 2016; Fleurbaey & Abi-Rafeh, 2016; Nurmi & Ahtiainen, 2018; U.S. Office of Management and Budget, 2023). Distributionally weighted BCA can be used to approximate *some* variants of the SWF methodology – but not all (Adler, 2016a, online supplement). Moreover, even for the variants that can be thus approximated, it is important to bear in mind that distributional weighting is an *approximation*: its assessment of policies can deviate from the SWF approach's, and will increasingly tend to do so for "large" policies, which involve significant changes in individuals' income, health, or other attributes relative to the status quo.

What this article addresses is the direct assessment of policy choice by the SWF methodology – not the proxying of that assessment via distributionally weighted BCA.

2. The SWF Framework

2.1. The SWF framework: a quick overview

The SWF framework, as described here,[1] is a methodology for implementing *welfarism.*[2] Welfarism is a family of ethical views that includes utilitarianism as its most prominent member, but reaches more broadly (Adler, 2012). Welfarists are *consequentialists.* Roughly speaking, consequentialists believe that the ethical status of an action depends upon what might happen were the action to be performed, and how good or bad these possible consequences might be. The philosophical notion of a "possible world" (a complete description of a possible history of the world) helps make this rough formulation more precise. At the core of any consequentialist ethical view is a world-ranking, which specifies for any pair of worlds d and d^* whether the first is better, worse, or equally good as the second – and does so in a well-behaved (transitive) fashion.

Welfarists take the world-ranking to be determined by individual well-being. This is captured in the Pareto axioms, Pareto Indifference and Strong Pareto. *Pareto Indifference*: If each person is equally well off in d^* as she is in d, then d^* and d are equally good. *Strong*

[1] By contrast, the literature in health economics that applies an SWF to some measure of individual health, cited in the Introduction, might be seen as non-welfarist – insofar as health is seen to be intrinsically important apart from well-being, rather than a proxy for well-being.

[2] Welfarism is, to be sure, controversial. See Adler (2012, chap. 1), reviewing the arguments for and against welfarism.

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

Pareto: If each person is at least as well off in d^* as she is in d, and at least one person is better off in d^* than d, then d^* is better than d.[3]

Consequentialists and, specifically, welfarists derive the ethical status of actions from the world-ranking. Let **P** be a set of choices (possible actions) facing a decision-maker. Welfarists aim to rank the choices in **P** as better or worse than each other, and to do so in light of the world ranking. But this is not straightforward. First, possible worlds are not cognitively tractable objects. A human decision-maker (even aided by computers) cannot write down a full description of even a single possible world, let alone each world in a set of worlds. Second, a human decision-maker, who is not omniscient, will not know for certain which world a given choice will lead to.

In short, welfarism needs a *decision procedure*: a methodology that is grounded in the world-ranking but implementable by human decision-makers, and that yields choice guidance. The SWF framework, in turn, is the most systematic such decision procedure. It is most appropriate for large-scale choices – in particular, governmental policies. I will therefore refer to **P** as a set of *policies.* A policy is some course of action that government might take: enacting a particular regulation, building infrastructure, disseminating information, etc. $\mathbf{P} = \{P, P^*, \ldots\}$, with P, P^*, etc. each a possible policy.

The SWF framework has the following components: a model population $\mathbf{I} = \{1, \ldots, N\}$; a set of outcomes $\mathbf{O} = \{x, y, \ldots\}$; a well-being measure $w(\cdot)$: the SWF proper, abbreviated as $\succcurlyeq$, which is a ranking of well-being vectors; the set **P** of policies, each represented as a probability distribution over outcomes; and an "uncertainty module" for the SWF, which is a rule for arriving at a ranking $\succcurlyeq^{\mathbf{P}}$ of the policy set in light of the well-being measure and the SWF. I will briefly discuss each component in turn. See Adler (2019, 2022) for a much fuller presentation of the SWF framework. These works, in turn, build upon a long-standing literature in welfare economics (Bossert & Weymark, 2004; Weymark, 2016).

I is a representation of the population of interest: those individuals whose well-being the decision-maker takes into account in choosing among policies. $\mathbf{I} = \{1, 2, \ldots, N\}$ is a set of N numbers, each denoting an individual. For simplicity, I will present the SWF framework using a fixed-population setup: each individual exists in all of the outcomes.[4]

An outcome x is a simplified and cognitively tractable model of a possible world. An outcome describes *some* of the features of worlds that are relevant to individual well-being. More specifically, each outcome x is a list of attribute bundles, one for each member of the population. $x = (b_1(x), \ldots, b_i(x), \ldots b_N(x))$. A bundle describes *some* of the types of individual attributes that determine well-being: attributes such as income, health, environmental quality, and leisure.

Bundles are *lifetime* bundles: individual i's lifetime well-being in x is determined by her bundle there.[5] The most simplified application of the SWF methodology employs

[3] A further, more technical aspect of welfarism is that the world-ranking corresponds to a *single* ranking of well-being vectors, rather than having "profile-dependent" rankings that vary depending on which well-being measure is used to map worlds onto well-being vectors. See Weymark (2016) and Adler (2019, 260–262). This aspect of welfarism is reflected in the SWF framework, see immediately below: $\succcurlyeq$ is a single vector-ranking rule rather than being indexed to the well-being measure $w(\cdot)$.

[4] See Adler (2019, 237–248; 2022, 102–106) for a brief discussion of how the SWF framework can be extended to the variable-population context. Blackorby *et al.* (2005) is the definitive work on this topic. Adler (forthcoming, chap. 8), discusses how the risk-and-attribute-profile apparatus described in Section 3 of this article can be extended to the variable-population context.

[5] See Adler (2012, chap. 6) for a defense of the proposition that welfarism should be specified in terms of lifetime, not sublifetime, well-being.

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

one-period bundles to model lifetime well-being: each individual exists for a single period in a given outcome, with his lifetime well-being determined by his attributes during this single period. A more complicated implementation employs *multi-period* bundles. Each bundle describes the individual's longevity (the number of periods that she is alive) and, for each period alive, her period attributes (income, health, leisure, etc.). The multi-period setup provides insights into tradeoffs between longevity and other well-being determinants, and will be the fulcrum for my analysis of health policy in Section 3.

The well-being measure $w(\cdot)$ maps bundles onto well-being numbers. These numbers represent *admissible* well-being comparisons, according to some theory of well-being. "Admissible" comparisons are those that the theory allows – that it sees as meaningful. Imagine that the theory allows for intra- and interpersonal comparisons of well-being levels and differences. Then $w(\cdot)$ will mirror such comparisons, as follows. (1) *Intrapersonal level comparisons.* $w(b_i) \geq w(b_i^*)$ — b_i and b_i^* two possible bundles for the same individual (i) – iff[6] i is at least as well off with b_i as b_i^*. (2) *Intrapersonal difference comparisons.* $w(b_i) - w(b_i^*) \geq w(b_i^{**}) - w(b_i^{***})$ — b_i, b_i^*, b_i^{**}, b_i^{***} four possible bundles for the same individual (i) – iff the difference in i's well-being between having b_i and having b_i^* is at least as large as the difference in i's well-being between having $b_i^{**}$and having b_i^{***}. (3) *Interpersonal level comparisons.*$w(b_i) \geq w(b_j)$ — b_i and b_j two possible bundles for different individuals (i and j, $i \neq j$) — iff i with b_i is at least as well off as j with b_j. (4) *Interpersonal difference comparisons.* $w(b_i) - w(b_j) \geq w(b_k) - w(b_l)$ — b_i, b_j, b_k, b_l four possible bundles for individuals who are not all identical (i, j, k, l, and not $i = j = k = l$) – iff the difference in well-being between i's having b_i and j's having b_j is at least as large as the difference in well-being between k's having b_k and l's having b_j.

The philosophical literature on well-being recognizes a wide range of types of well-being theories (see Adler, 2019, chap. 3, and sources cited therein). These can be grouped into three families: "experientialist" theories, according to which an individual's well-being is reducible to the mental states that he experiences (pains, pleasures, feelings of happiness, etc.); preference-based theories, according to which an individual's well-being depends on the extent to which her preferences are satisfied; and objective-good theories, according to which well-being consists in the realization of various goods that are not reducible to either experiences or preference-satisfaction. These three theories correspond to different strands in economics: experientialism, to the burgeoning field of happiness economics; preference-based theories, to the view of well-being in neoclassical economics; and objective-good theories, to the literature on "capabilities."

The SWF framework is agnostic about the nature of well-being. Any well-being theory can be plugged into the framework – with one big caveat. The theory must allow for interpersonal well-being comparisons. The SWF framework founders if combined with a well-being theory that eschews interpersonal comparisons – that counts as admissible *only* intrapersonal well-being level comparisons, or *only* intrapersonal well-being level and difference comparisons. I will explain why in a moment.

A well-being vector (denoted here with a bold lower-case letter such as **w** or **v**) is a list of well-being numbers, one for each individual in the population. $\mathbf{w} = (\mathbf{w}_1, \ldots, \mathbf{w}_i, \ldots, \mathbf{w}_N)$, with $\mathbf{w}_i$ the well-being number of individual i. The well-being measure $w(\cdot)$ converts a given

[6] "Iff" is shorthand for "if and only if."

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

outcome x into a well-being vector. With $\mathbf{w}(x)$ the well-being vector corresponding to x, we have $\mathbf{w}(x) = (w(b_1(x)), \ldots, w(b_i(x)), \ldots w(b_N(x)))$.

The SWF proper is a rule for ranking well-being vectors. Many such rules are possible, but four predominate in the literature: the utilitarian SWF, the prioritarian family of SWFs, the leximin SWF, and the rank-weighted family of SWFs. (In what follows, "$\mathbf{w} \succcurlyeq \mathbf{v}$" means that $\mathbf{w}$ is ranked by the SWF at least as good as $\mathbf{v}$; "$\mathbf{w} \succ \mathbf{v}$," that $\mathbf{w}$ is ranked better than $\mathbf{v}$; and "$\mathbf{w} \sim \mathbf{v}$," that the two are ranked equally good.)

The utilitarian SWF: $\mathbf{w} \succcurlyeq \mathbf{v}$ iff $\sum_{i=1}^{N} \mathbf{w}_i \geq \sum_{i=1}^{N} \mathbf{v}_i$.

A prioritarian SWF: Each such SWF is defined by a strictly increasing, strictly concave and continuous function $g(\cdot)$, the "transformation function." $\mathbf{w} \succcurlyeq \mathbf{v}$ iff $\sum_{i=1}^{N} g(\mathbf{w}_i) \geq \sum_{i=1}^{N} g(\mathbf{v}_i)$.

The leximin SWF: Let $\widehat{\mathbf{w}}$ be a vector rearranging the elements of $\mathbf{w}$ from smallest to largest. Then: (1) $\mathbf{w} \sim \mathbf{v}$ iff $\mathbf{v}$ is a permutation of $\mathbf{w}$; and (2) $\mathbf{w} \succ \mathbf{v}$ iff there is $j \leq N$ such that $\widehat{\mathbf{w}}_i = \widehat{\mathbf{v}}_i$ for all $i < j$ and $\widehat{\mathbf{w}}_j > \widehat{\mathbf{v}}_j$.

The rank-weighted family of SWFs: Each such SWF is defined by a list of N positive and strictly decreasing weights: $k_1, k_2, \ldots, k_N$ such that $k_1 > k_2 > \ldots > k_N$. $\mathbf{w} \succcurlyeq \mathbf{v}$ iff $\sum_{i=1}^{N} k_i \widehat{\mathbf{w}}_i \geq \sum_{i=1}^{N} k_i \widehat{\mathbf{v}}_i$.

An SWF's vector ranking (whatever it may be) immediately generates a ranking of the outcome set: outcome x at least as good as outcome y iff $\mathbf{w}(x) \succcurlyeq \mathbf{w}(y)$.

The choice between these different types of SWFs is an *ethical* choice. The question how to compare possible arrangements of well-being among the population is a matter for ethical debate. In a given legal system, the legal authority to make this ethical choice – to select the SWF methodology as opposed to a different approach (e.g. BCA), and if so to use one SWF rather than others – will be vested in certain governmental officials or institutions (e.g. an elected President, a legislature).

"Prioritarianism" has that name because it gives *priority* to well-being changes affecting worse-off individuals, as can be seen in Figure 1. The degree of such priority depends upon the concavity of the transformation function. I have argued at length in favor of prioritarianism, as against utilitarianism, rank-weighted SWFs, and leximin (Adler, 2012; Adler & Holtug, 2019).

The key difference between utilitarian and prioritarian SWFs concerns the *Pigou–Dalton* axiom. Pigou–Dalton is an *equity* axiom. It captures, formally, whether the SWF is sensitive to the distribution of well-being.

Pigou–Dalton: Let $\mathbf{v}$ be reached from $\mathbf{w}$ via a pure, gap-diminishing transfer of well-being from a better-off to a worse-off person, affecting no one else.[7] Then $\mathbf{v} \succ \mathbf{w}$.

Every prioritarian SWF satisfies the Pigou–Dalton axiom. The utilitarian SWF does not; a pure transfer of well-being from a better- to a worse-off person does not change the sum total of well-being.

[7] That is, $\mathbf{w}_i > \mathbf{w}_j$; $\mathbf{v}_i = \mathbf{w}_i - \Delta w$, $\mathbf{v}_j = \mathbf{w}_j + \Delta w$, with $\Delta w > 0$; $|\mathbf{v}_i - \mathbf{v}_j| < |\mathbf{w}_i - \mathbf{w}_j|$; and $\mathbf{w}_k = \mathbf{v}_k$ for all $k \neq i, j$.

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

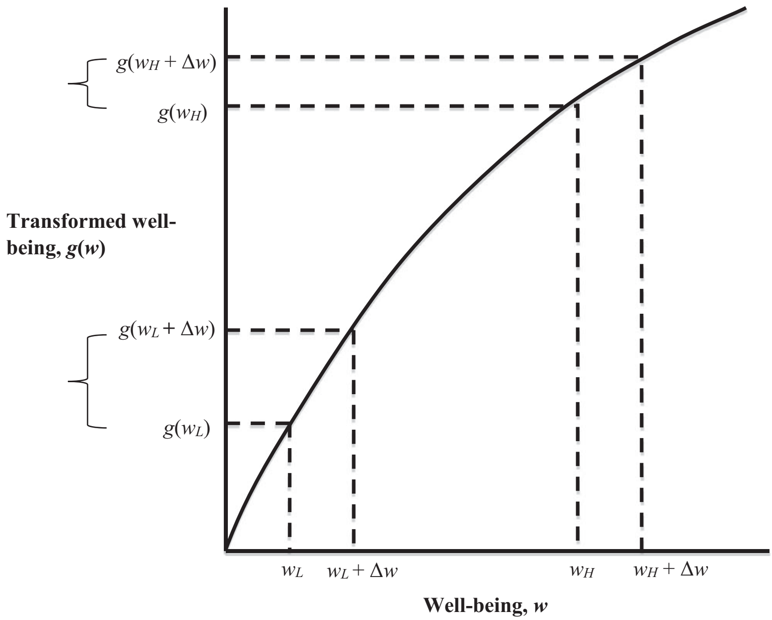

Figure 1. A prioritarian transformation function. *Explanation*: Let w_L be the well-being level of a worse-off person, and let w_H be the well-being level of a better-off person. An increment Δw to the worse-off person's well-being produces a bigger change in transformed well-being (as seen on the *y*-axis) than the same increment to the better-off person's well-being.

The leximin SWF and rank-weighted SWFs also satisfy the Pigou–Dalton axiom, but they have significant disadvantages as compared to prioritarian SWFs – or so I have argued. The leximin SWF is absolutist. Consider a *leaky* transfer of well-being between a transferor (the individual initially better off) and a transferee (the individual initially worse off) which is such that the transferor ends up better off than where the transferee started; no one else is affected. The transfer is "leaky" in that the transferor's well-being is reduced by Δw, while the transferee's increases by some fraction of the loss to the transferor – by $\rho \Delta w$, with ρ between 0 and 1. Leximin sees every such transfer as an ethical improvement, regardless of how small ρ is. By contrast, every prioritarian SWF will favor some such transfers (if ρ is sufficiently large) but disapprove others (if ρ is too small).[8]

Prioritarianism can be seen as filling the "space" of distribution sensitivity between utilitarianism, at one extreme, and leximin, at the other. Utilitarianism is wholly insensitive to equity: even a pure transfer of well-being from a better- to a worse-off person (affecting no one else), which shrinks the gap between them at *no* cost to overall well-being, is a matter of indifference to utilitarianism. Leximin is wholly insensitive to overall well-being. Any transfer from a better- to a worse-off individual (affecting no one else) that leaves the

[8] Let *i* be the transferor and *j* the transferee, with $\mathbf{w}_i$ and $\mathbf{w}_j$ their respective starting-point levels. With $\Delta w > 0$ and $(\mathbf{w}_i - \Delta w) > \mathbf{w}_j$, there is a unique β, $0 < \beta < 1$, such that $g(\mathbf{w}_i) - g(\mathbf{w}_i - \Delta w) = g(\mathbf{w}_j + \beta \Delta w) - g(\mathbf{w}_j)$. β depends on the transformation function $g(\cdot)$ and on $\mathbf{w}_i$ and $\mathbf{w}_j$. The leaky transfer is an ethical improvement/worsening/a matter of ethical indifference if ρ is greater than/less than/equal to β.

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

transferor better off than where the transferee started, is favored by leximin – regardless of the degree of leakage and the loss to overall well-being. Prioritarianism balances overall well-being against equity, and offers the decision-maker flexibility in doing so. As $g(\cdot)$ becomes increasingly concave, less weight is given to overall well-being: prioritarianism approaches leximin. As $g(\cdot)$ becomes less concave, and approaches linearity, less weight is given to equity; prioritarianism approaches utilitarianism.

Rank-weighted SWFs are less tractable than prioritarian SWFs. Consider once again the case of a leaky transfer that leaves the transferor better off than where the transferee started, and affects no one else. For prioritarians, knowing the starting and ending well-being levels of the two individuals, together with the transformation function $g(\cdot)$, is sufficient to determine whether the transfer is an ethical improvement. For those who endorse a rank-weighed SWF, this is not sufficient information. Whether the transfer is an ethical improvement will generally depend upon how the starting and ending well-being levels of transferee and transferor compare to the well-being levels of everyone else – even though they are unaffected.[9]

Formally, prioritarian SWFs (as well as the utilitarian SWF and leximin SWF), by contrast with rank-weighted SWFs, satisfy an axiom of *Separability*:

> *Separability*: Assume that some subset of the population $\mathbf{M}$ is such that $\mathbf{w}_j = \mathbf{v}_j$ for each j in $\mathbf{M}$. Then the ranking of $\mathbf{w}$ as compared to $\mathbf{v}$ is independent of what these well-being levels are.[10]

The tractability benefits of Separability will come into clearer view in Section 2.2, when we consider the topic of uncertainty modules.

Let us turn now to the question of interpersonal comparability. A basic axiom of the SWF framework, which I have referred to as the "Fundamental Principle of Invariance" (Adler, 2019, chap. 2), says: the SWF's ranking of outcomes should be invariant to replacement of well-being measure $w(\cdot)$ by an *informationally equivalent* well-being measure $w^+(\cdot)$. If $w(\cdot)$ accurately represents all of the admissible well-being comparisons, according to whichever well-being theory is being used, and $w^+(\cdot)$ does so as well, then the SWF's outcome ranking should be the same whether $w(\cdot)$ is used to map outcomes onto vectors or instead $w^+(\cdot)$ is. For the SWF to rank outcomes one way with $w(\cdot)$, and a different way with $w^+(\cdot)$, would mean that the outcome ranking is *arbitrary* from the perspective of welfarism: the ranking depends on the choice between the two measures, a choice that cannot be justified on welfarist grounds because the measures contain the very same well-being information.

Assume now that our well-being theory sees only intrapersonal comparisons as admissible: either nothing more than intrapersonal level comparisons, or nothing more than intrapersonal level and difference comparisons. It permits neither interpersonal level comparisons, nor interpersonal difference comparisons. Note that if well-being measure $w(\cdot)$ accurately represents the theory's intrapersonal comparisons, then so does any other

[9] Whether the transfer is an ethical improvement will depend upon the weights that are applied to the transferor's and transferee's well-being levels in the starting and ending points – which in turn depend upon where those well-being levels are located in the population distribution of well-being.

[10] Here is a more precise statement. Let $\mathbf{M}$ be any subset of $\mathbf{I}$, and let $\mathbf{M}^+ = \mathbf{I} \setminus \mathbf{M}$ (all individuals not in $\mathbf{M}$). Assume that $\mathbf{w}$, $\mathbf{v}$, $\mathbf{w}^*$, and $\mathbf{v}^*$ are as follows. For all i in $\mathbf{M}$, $\mathbf{w}_i = \mathbf{v}_i$ and $\mathbf{w}_i{}^* = \mathbf{v}_i{}^*$. For all j in $\mathbf{M}^+$, $\mathbf{w}_j = \mathbf{w}_j{}^*$ and $\mathbf{v}_j = \mathbf{v}_j{}^*$. Then $\mathbf{w} \succcurlyeq \mathbf{v}$ iff $\mathbf{w}^* \succcurlyeq \mathbf{v}^*$.

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

well-being measure $w^{+}(\cdot)$ that is an individual-specific cardinal rescaling of $w(\cdot)$.[11] The two measures are informationally equivalent *if* interpersonal comparisons are inadmissible.

However, as Table 1 illustrates, none of the major types of SWFs are invariant to individual-specific cardinal rescalings of the well-being measure. For each such SWF, swapping one well-being measure $w(\cdot)$ for another measure $w^{+}(\cdot)$ that makes exactly the same intrapersonal comparisons of levels and differences leads to a change in the outcome ranking. In short, *if* interpersonal comparisons are inadmissible, then all these SWFs violate the Fundamental Principle of Invariance. Indeed, it can be shown that *any* non-dictatorial SWF will violate the Fundamental Principle of Invariance if interpersonal comparisons are inadmissible.[12]

Should we, in fact, endorse a well-being theory that rejects interpersonal comparisons? Some economists will say "yes" and therefore reject the SWF framework. But interpersonal comparisons are a matter of common sense (Adler, 2022, 75–80). To be sure, *how* to make interpersonal well-being comparisons for purposes of the SWF framework implicates contested ethical questions[13] – but the decision to adopt *any* policy-analysis methodology implicates contested ethical questions.

Special puzzles *do* arise in explaining how a *preference-based* well-being theory allows for interpersonal comparisons. Adler (2016b, 2019) shows how to construct a well-being measure that respects individual preferences and makes both intrapersonal and interpersonal comparisons of levels and differences, by using individuals' von Neumann/Morgenstern utility functions. Another fairly widespread methodology for preference-based interpersonal comparisons is the so-called "equivalence approach" (as implemented, e.g., in the use of "equivalent income" as the measure of an individual's well-being given her preferences) (Adler & Decancq, 2022; Cookson *et al.,* 2022; Fleurbaey, 2016).

2.2. The uncertainty module

In what follows, I focus on utilitarianism (the most influential version of welfarism, historically and up through the present) and prioritarianism (which I take to be utilitarianism's strongest competitor).

The SWF framework uses an "uncertainty module" to capture the decision-maker's uncertainty about which outcome would result were a given policy to be implemented. Each policy in $\mathbf{P}$ is associated with a probability distribution over $\mathbf{O}$, the set of outcomes. $\pi_P(x)$ is the probability of outcome x, were policy P to be chosen.[14] $\succcurlyeq^{\mathbf{P}}$, recall, is the ranking of the policy set. An *uncertainty module* for a particular SWF ($\succcurlyeq$, a ranking of well-being vectors) is a rule for arriving at $\succcurlyeq^{\mathbf{P}}$ in light of the probabilities over outcomes for each P; the well-being measure $w(\cdot)$; and the SWF. Each SWF has multiple uncertainty modules. Every such

[11] $w^{+}(\cdot)$ is an individual-specific cardinal rescaling of $w(\cdot)$ if there exist constants $c_i > 0$, d_i for each individual i such that: for every bundle b_i (b_i some bundle held by i), $w^{+}(b_i) = c_i w(b_i) + d_i$. It is easy to see that $w^{+}(\cdot)$ make the very same intrapersonal level and difference comparisons as $w(\cdot)$. For each individual i, and all bundles b_i, ${b_i}^{*}$, ${b_i}^{**}$, and ${b_i}^{***}$: $w(b_i) \geq w({b_i}^{*})$ iff $w^{+}(b_i) \geq w^{+}({b_i}^{*})$; and $w(b_i) - w({b_i}^{*}) \geq w({b_i}^{**}) - w({b_i}^{***})$ iff $w^{+}(b_i) - w^{+}({b_i}^{*}) \geq w^{+}({b_i}^{**}) - w^{+}({b_i}^{***})$.

[12] See Adler (2019, 45) for a more precise statement of this result.

[13] A variety of well-being theories, all of which admit such comparisons, might be adopted; which theory to adopt is an ethical question.

[14] The presentation here assumes that the distribution has finite support; for each P, only a finite number of outcomes x are such that $\pi_P(x) > 0$.

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

Table 1. *An individual-specific cardinal rescaling of the well-being measure*

	Original well-being numbers				Individual-specific cardinal rescaling				
								Scaling factors	
	Outcome x	y	z		Outcome x	y	z	c_i	d_i
Abel	9	4	21	Abel	9	4	21	1	0
Bob	25	25	21	Bob	230	230	190	10	–20
Chloe	49	16	21	Chloe	7.9	4.6	5.1	0.1	3
Diana	1	36	21	Diana	400	14,400	8,400	400	0
Util. score	84	81	84		646.9	14,638.6	8,616.1		
Prior. score($\sqrt{}$)	16	17	18.3		41	139.3	112.3		
Rk–wtd score (integer)	130	150	210		918.6	14,889.8	8,863.4		
Leximin	$z \succ y \succ x$				$x \succ z \succ y$				

Explanation: The left of the table displays well-being vectors assigned to outcomes x, y, and z by a well-being measure $w(\cdot)$, while the right displays vectors per an individual-specific cardinal rescaling of $w(\cdot)$, $w^+(\cdot)$ such that $w^+(b_i) = c_i w(b_i) + d_i$, $c_i > 0$. The table illustrates that neither the utilitarian SWF, nor a prioritarian SWF, nor a rank-weighted SWF, nor the leximin SWF are invariant to an individual-specific cardinal rescaling of well-being numbers. The square root transformation function is used for the prioritarian SWF, and the rank-weighted SWF uses integer weights $(k_1 = N, k_2 = N-1, \ldots, k_N = 1)$. *Source*: Adler (2022, 78).

module must satisfy a consistency constraint, which applies to the ranking of "degenerate" policies (policies that assign probability 1 to one outcome and 0 to all others): if P leads with probability 1 to outcome x, and P^* with probability 1 to outcome x^*, then the policies must be ranked according to the SWF's ranking of the well-being vectors associated with x and x^*. Once nondegenerate policies come into the picture, however, the different modules for a given SWF may diverge.

In the literature on SWFs under uncertainty (for overviews, see Adler, 2012 chap. 7, 2019, 2022; Mongin & Pivato, 2016), the dominant uncertainty module for the utilitarian SWF is a rule that I will term "simple utilitarianism" (SU). How to apply prioritarianism under uncertainty is more contested. Three modules predominate: "ex post prioritarianism" (EPP), "ex ante prioritarianism" (EAP), and expected equally-distributed-equivalent prioritarianism (EEDEP).[15] Each of these assigns numerical scores to policies and ranks them according to the scores.[16] I will denote the scores as, respectively, $S^{SU}(P)$, $S^{EPP}(P)$, $S^{EAP}(P)$, and $S^{EEDEP}(P)$; "S" indicates "social welfare" and the superscript the particular module at hand. The scores are calculated as follows. ($g(\cdot)$, as above, is the transformation function that defines a specific prioritarian SWF; EPP, EAP, and EEDEP are all modules for that SWF.)

$$\textit{Simple Utilitarianism: } S^{SU}(P) = \sum_x \pi_P(x) \sum_{i=1}^{N} \mathbf{w}_i(x) = \sum_{i=1}^{N} \sum_x \pi_P(x) \mathbf{w}_i(x).$$

[15] The expected equally distributed equivalent was pioneered by Fleurbaey (2010), a major contribution to the literature on SWFs under uncertainty.

[16] $S(\cdot)$ is a real-valued function, with $S(P)$ the number ("score") assigned to policy P. Each of these modules employs some such $S(\cdot)$ and keys $\succcurlyeq^{\mathbf{P}}$ to it: $P \succcurlyeq^{\mathbf{P}} P^*$ iff $S(P) \geq S(P^*)$.

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

Ex Post Prioritarianism: $S^{EPP}(P) = \sum_x \pi_p(x) \sum_{i=1}^{N} g(\mathbf{w}_i(x)) = \sum_{i=1}^{N} \sum_x \pi_P(x) g(\mathbf{w}_i(x))$.

Ex Ante Prioritarianism: $S^{EAP}(P) = \sum_{i=1}^{N} g\left(\sum_x \pi_P(x)\mathbf{w}_i(x)\right)$.

Expected Equally Distributed Equivalent Prioritarianism: $S^{EEDEP}(P) = \sum_x \pi_P(x) g^{-1}\left(\sum_{i=1}^{N} g(\mathbf{w}_i(x))/N\right)$.

In a nutshell: The SU score is the expected sum of individuals' well-being or, equivalently, the sum across individuals of expected well-being. The EPP score is the expected sum of individuals' *transformed* well-being or, equivalently, the sum across individuals of expected transformed well-being. The EAP score is the sum, across individuals, of transformed expected well-being. The prioritarian "equally distributed equivalent" for a given well-being vector **w** is that well-being level *w* such that a vector with everyone at *w* is ranked equally good by the prioritarian SWF as **w**. The EEDEP score is the expected value of these equally distributed equivalents.

Uncertainty axioms are constraints that, it might be proposed, the uncertainty module should satisfy. There are a range of such axioms that seem ethically plausible, but I will focus on three: *Ex Ante Pareto*, *Dominance*, and *Policy Separability.*[17] These axioms will help clarify what is at stake in the choice between EPP, EAP, and EEDEP, and why a single dominant module has emerged for utilitarianism (SU) but not so for prioritarianism. And *Policy Separability* will be the foundation for the analysis of health policy in Section 3.

> *Ex Ante Pareto.* (1) *Ex Ante Pareto Indifference.* If each person's expected well-being with P is equal to her expected well-being with P^*, the policies are equally good. (2) *Ex Ante Strong Pareto.* If at least one person has greater expected well-being with P than P^*, and no one has lower expected well-being, P is a better policy.
>
> *Dominance.* If every outcome with non-zero probability given P has a well-being vector that is preferred by the SWF to the well-being vector of every outcome with non-zero probability given P^*, the uncertainty module for that SWF should prefer P to P^*.
>
> *Policy Separability.* (1) If each person faces the same lottery over well-being with P as she does with P^*, the two policies are equally good. (2) Let P and P^* be such that some individuals face different well-being lotteries with the two policies, while other individuals each face the same well-being lottery with P as he does with P^*. Then the P/P^* ranking is invariant to which well-being lottery each person in the latter group faces.[18]

Table 2 states how the various modules fare with respect to these three uncertainty axioms.

Table 3 displays a basic dilemma with respect to a prioritarian module. No such module can satisfy both Ex Ante Pareto and Dominance. EPP and EEDEP satisfy Dominance, at the inevitable cost of violating Ex Ante Pareto; EAP satisfies Ex Ante Pareto, at the inevitable cost of violating Dominance.

[17] See Appendix A for a formal statement of these axioms.

[18] The second prong of Policy Separability implies the first. See Appendix A.

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

Table 2. *Uncertainty modules and axioms*

	Ex Ante Pareto	Dominance	Policy Separability
SU	Yes	Yes	Yes
EPP	No	Yes	Yes
EAP	Yes	No	Yes
EEDEP	No in general; yes if individuals are identically situated (see note 23 for explanation)	Yes	No

By contrast, as Table 3 also illustrates, utilitarians face no such dilemma. SU satisfies both Ex Ante Pareto and Dominance.

In order to grasp the difference between utilitarianism and prioritarianism with respect to Ex Ante Pareto and Dominance, it is critical to understand that what Dominance demands *depends upon the SWF.* Assume that P has probability 0.5 of yielding well-being vector $\mathbf{w} = (25, 81)$ and probability 0.5 of yielding well-being vector $\mathbf{w}^* = (81, 25)$, while policy P^* has probability 1 of yielding vector $\mathbf{v}$ = (50, 50). Consider the prioritarian SWF using a square-root transformation function. Dominance requires that the uncertainty module for this SWF prefer P^* to P; according to *this* SWF, $\mathbf{v}$ is better than both $\mathbf{w}$ and $\mathbf{w}^*$. By contrast, Dominance requires that the uncertainty module for the *utilitarian* SWF prefer P to P^*; according to *that* SWF, both $\mathbf{w}$ and $\mathbf{w}^*$ are better than $\mathbf{v}$.[19]

It is also important to stress that the prioritarian SWF *satisfies* the Pareto axiom at the level of well-being vectors.[20] Problems emerge only with uncertainty. Finally, it should be observed that the conflict between Dominance and Ex Ante Pareto transcends prioritarianism. It can be shown that *any* SWF which satisfies Pigou–Dalton faces this conflict.[21]

In my own view, the Dominance axiom is compelling (if one policy is certain to yield a better outcome than a second policy, then surely the first policy is better), as are the Pareto axioms at the level of well-being vectors, while the *Ex Ante* Pareto axioms are less persuasive. For example, in the kind of case illustrated by the top half of Table 3, Ex Ante Strong Pareto requires a preference for one policy over a second even though the decision-maker can be sure that, if the individuals were fully informed, at least one would prefer the second policy. But how to handle conflicts between Dominance and Ex Ante Pareto is ethically contestable. Some readers may believe that the inevitability of such conflict given a concern for equity (Pigou–Dalton) is a powerful argument for the utilitarian SWF. Others may reject this point of view and endorse prioritarianism, but with the caveat that it should be

[19] $\sqrt{25} + \sqrt{81} < \sqrt{50} + \sqrt{50}$; but $25 + 81 > 50 + 50$.

[20] If $\mathbf{w}_i = \mathbf{v}_i$ for all i, $\mathbf{w} \sim \mathbf{v}$ (Pareto Indifference for well-being vectors; satisfied by *any* SWF); and if $\mathbf{w}_i \geq \mathbf{v}_i$ for all i and $\mathbf{w}_j > \mathbf{v}_j$ for at least one j, $\mathbf{w} \succ \mathbf{v}$ (Strong Pareto for well-being vectors; satisfied by prioritarian SWFs as well as the utilitarian SWF, leximin SWF, and rank-weighted SWFs).

[21] More precisely, if an SWF satisfies Pigou–Dalton, it cannot be applied under uncertainty in a manner that satisfies both Dominance and Ex Ante Pareto Indifference; and if it satisfies Pigou–Dalton and is minimally leak tolerant (there is some gap-diminishing leaky transfer from a better- to a worse-off person, leaving everyone else unaffected, that the SWF prefers), then it cannot be applied under uncertainty in a manner that satisfies both Dominance and Ex Ante Strong Pareto. See Adler (2019, 140–144).

Table 3. *Dominance and Ex Ante Pareto*

	Policy P			Policy P^{+}		
	$\underline{\pi = 0.5}$	$\underline{\pi = 0.5}$	Expected well-being	$\underline{\pi = 0.5}$	$\underline{\pi = 0.5}$	Expected well-being
Lillian	70	30	50	$50-\varepsilon$	$50-\varepsilon$	$50-\varepsilon$
Maya	30	70	50	$50-\varepsilon$	$50-\varepsilon$	$50-\varepsilon$
	Policy P^{*}			Policy P^{**}		
	$\underline{\pi = 0.5}$	$\underline{\pi = 0.5}$	Expected well–being	$\underline{\pi = 0.5}$	$\underline{\pi = 0.5}$	Expected well–being
Lillian	70	30	50	50	50	50
Maya	30	70	50	50	50	50

Explanation: Each of the policies (P, P^{+}, P^{*}, and P^{**}) leads to some outcome with probability 0.5 and some other outcome with probability 0.5. The table displays the well-being vectors corresponding to the outcomes.
In the top half of the table, for any prioritarian SWF, there is some cutoff value $c > 0$ (which depends on the transformation function) such that the well-being vector $(50-\varepsilon, 50-\varepsilon)$ is preferred by the SWF to (70, 30) and (30, 70) for every ε, $0 < \varepsilon < c$. Dominance requires that the module for that SWF rank P^{+} over P. But note that Ex Ante Strong Pareto requires that P be ranked above P^{+}. Dominance requires that the module for the utilitarian SWF rank P over P^{+}, since the well-being vectors (70, 30) and (30, 70) are preferred by the utilitarian SWF to $(50-\varepsilon, 50-\varepsilon)$.
SU ranks P over P^{+}, as does EAP. EPP and EEDEP rank P^{+} over P.
In the bottom half of the table, Dominance requires that the module for any prioritarian SWF rank P^{**} over P^{*}, since any prioritarian SWF prefers the well-being vector (50, 50) to (70, 30) and (30, 70). However, Ex Ante Pareto Indifference requires that P^{**} and P^{*} be ranked equally good. Dominance does not here constrain the module for a utilitarian SWF, since that SWF is indifferent between (50, 50) and both (70, 30) and (30, 70).
SU ranks P^{*} and P^{**} equally good, as does EAP. EPP and EEDEP rank P^{**} over P^{*}.

applied under uncertainty with the EAP module (so as to satisfy the Ex Ante Pareto axiom, at the cost of Dominance).

Let us turn now to Policy Separability. Policy Separability is a *tractability* axiom. Note that each policy is associated with a list of bundle lotteries, one for each person in the population. Let us say that an individual is “unaffected” if she faces the same lottery over bundles for each policy in **P**. If Policy Separability holds true of an uncertainty module, knowing how each policy in **P** maps onto a list of bundle lotteries, one for each *affected* person, is sufficient information for ranking the policy set. Policies need not be characterized as probability distributions over whole outcomes. Instead, it is sufficient to ascertain how policies endow individuals with bundle lotteries – and not *all* individuals, just the affected ones. In other words, information about the *joint* probability distribution of bundles among all N individuals in the population, or among the affected group, is not needed to rank policies. See Table 4, illustrating Policy Separability.

The relation between Separability (an axiom regarding the ranking of well-being vectors) and Policy Separability (an uncertainty axiom) is subtle. If an SWF fails Separability, none of its modules will satisfy Policy Separability.[22] If an SWF satisfies Separability, some of its modules will satisfy Policy Separability, while others will not. This is illustrated by

[22] As mentioned earlier, the module’s ranking of “degenerate” policies (those that assign probability 1 to some outcome and 0 to all others) should conform to the SWF’s ranking of well-being vectors. Since the latter fails Separability, the module’s ranking of these degenerate policies, at the very least, cannot satisfy Policy Separability.

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

Table 4. *Policy Separability*

	Policy Separability: Prong 1			
	Policy P		Policy P^+	
	$\pi = 0.5$	$\pi = 0.5$	$\pi = 0.5$	$\pi = 0.5$
Juan	b	b^*	b^*	b
Mitch	b'	b''	b'	b''
	Policy Separability: Prong 2			
	Policy P^*		Policy P^{**}	
	$\pi = 0.5$	$\pi = 0.5$	$\pi = 0.5$	$\pi = 0.5$
Amaya	b	b^*	b^{**}	b^{***}
Frank	b'	b''	b'''	b''''
James	b^+	b^{++}	b^+	b^{++}
	PolicyP'		PolicyP''	
	$\pi = 0.5$	$\pi = 0.5$	$\pi = 0.5$	$\pi = 0.5$
Amaya	b	b^*	b^{**}	b^{***}
Frank	b'	b''	b'''	b''''
James	b^{+++}	b^{++++}	b^{+++}	b^{++++}

Explanation: Each of the policies ($P, P^+, P^*, P^{**}, P', P''$) leads to some outcome with probability 0.5 and some other outcome with probability 0.5. The table displays the bundles that each person in the population receives in these outcomes.
The top part of the table illustrates Prong 1 of Policy Separability. Note that the probability distribution of outcomes for policy P is different from that for policy P^+. However each individual in the population faces the same bundle lottery with P as he does with P^+: Juan a lottery with equal chances of b and b^*, Mitch a lottery with equal chances of b' and b''. Therefore, each faces the same lottery over well-being levels with P as he does with P^+. Prong 1 of Policy Separability thus requires the two policies to be ranked equally good.
The bottom part of the table illustrates Prong 2 of Policy Separability. James is unaffected in the P^*/P^{**} comparison: he faces the same bundle lottery with the two policies. Therefore, he faces the same well-being lottery with the two policies. Prong 2 of Policy Separability requires that the P^*/P^{**} ranking be invariant to *which* well-being lottery James faces. Note that P' is the same as P^*, and P'' the same as P^{**}, with respect to the bundle lotteries faced by Amaya and by Frank. The only difference is that James faces a different bundle lottery with P' and P'' than he does with P^* and P^{**}. Thus, Prong 2 of Policy Separability requires that P'/P'' ranking be the same as the P^*/P^{**} ranking.

prioritarianism. The prioritarian SWF satisfies Separability; its EPP and EAP modules satisfy Policy Separability; its EEDEP module does not.

EEDEP satisfies Ex Ante Pareto when individuals are identically situated[23] although not in general (see Table 3) or even when some are identically situated and everyone else is unaffected;[24] it purchases this limited compliance with Ex Ante Pareto at the cost of violating Policy Separability. EPP does not satisfy Ex Ante Pareto even when individuals are identically situated, but satisfies Policy Separability.[25] To my mind, EPP is on balance

[23] Meaning: If well-being is distributed perfectly equally in each of the outcomes assigned nonzero probability by P and by P^*, then (1) if Ex Ante Strong Pareto requires that P be ranked above P^*, EEDEP ranks P above P^*; and (2) if alternatively Ex Ante Pareto Indifference requires that P and P^* be ranked equally good, EEDEP does so.

[24] See Adler (2019, chap. 4), discussing "heartland cases."

[25] If a module satisfies Ex Ante Pareto, it satisfies the first prong of Policy Separability, but the converse is not true – as illustrated by EPP.

more attractive than EEDEP, given the large tractability benefits of Policy Separability, but that conclusion is certainly contestable.

3. SWFs and health policy

By "health policy," I mean a governmental intervention that seeks to improve individuals' health and/or reduce their fatality risks, typically at some cost in income or other non-health attributes. This is a broad definition of health policy, in that it encompasses policies designed to reduce fatality risks that result either from disease or from other sources (e.g. auto accidents).

As mentioned in the Introduction, a recent branch of the SWF literature employs the SWF framework to evaluate health policies – and does so by applying the SWF to a measure of *well-being*, not health alone. This measure is intended to reflect all the sources of well-being (up to the limits of modeling tractability), including material resources (income, consumption, and wealth).

What follows is an apparatus for applying the utilitarian SWF to health policies via the simple-utilitarian uncertainty module (SU), and for applying a prioritarian SWF via the ex post prioritarian module (EPP) or ex ante prioritarian module (EAP). This apparatus builds upon Adler (2017, 2019 chap. 5, 2020a, 2020b, Forthcoming) and Adler *et al.* (2021). For reasons that will become apparent momentarily, I will refer to it as the "risk-and-attribute-profile" apparatus. See Appendix C for a formal presentation.

Because SU, EPP, and EAP satisfy Policy Separability, the scores they assign to a given policy can be expressed as a function of the array of individual bundle lotteries associated with that policy. The scores, restated in this "bundle lottery" form, are as follows. $\rho_{P,i}(b)$ is the probability with policy P that individual i receives lifetime bundle b.

$$\textit{Simple Utilitarianism}: \ S^{SU}(P) = \sum_{i=1}^{N} \sum_{b} \rho_{P,i}(b) w(b).$$

$$\textit{Ex Post Prioritarianism}: \ S^{EPP}(P) = \sum_{i=1}^{N} \sum_{b} \rho_{P,i}(b) g(w(b)).$$

$$\textit{Ex Ante Prioritarianism}: \ S^{EAP}(P) = \sum_{i=1}^{N} g\left(\sum_{b} \rho_{P,i}(b) w(b) \right).$$

The risk-and-attribute-profile apparatus associates a given policy P with a lottery over lifetime bundles for each individual i, as follows. Individual lifetimes are divided into periods (e.g. years), with T periods the maximum possible lifespan. Calendar time is divided into past, present, and future. The present "age" of individual i – the number of periods that she has survived so far – is A_i. (Thus, the present time is at the beginning of period $A_i + 1$ for individual i.)[26]

[26] Note that Adler *et al.* (2021) numbers periods differently than the current article, setting the present period as period A_i for individual i rather than $(A_i + 1)$. This difference has no substantive impact: what matters for the risk-

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

Death is modeled as follows. An individual either dies at the beginning of period t of his life (in which case he has the attribute "Dead" during the period), or he survives to the end of the period (in which case he has some bundle of period attributes b^t).

Policy P endows individual i with a "risk profile" $\mathbf{p}_{P,i}$ and an "attribute profile" $\mathbf{b}_{P,i}$. The risk profile is a list of survival probabilities, starting with the current period and ending with the last possible period. A "survival probability" p_i^t is the probability that i survives to the end of period t, conditional on being alive at the beginning. That is, risk profile $\mathbf{p}_{P,i} = \left(p_{P,i}^{A_i+1}, \ldots, p_{P,i}^{T}\right)$, with $p_{P,i}^t$ individual i's survival probability for period t with policy P.

The attribute profile is a list of period bundles, starting with the first period of i's life and ending with the last possible period. $\mathbf{b}_{P,i} = \left(b_{P,i}^1, \ldots, b_{P,i}^T\right)$. $b_{P,i}^t$ is the period bundle that i will receive in period t, if she survives to the end rather dying earlier. Period bundles can be specified in terms of *any* attributes that the analyst takes to be relevant to well-being: income, health quality, leisure, and so forth. If the well-being measure is preference-based and the analyst wishes to take account of preference heterogeneity in the population, then period bundles are "hybrid bundles" – specifying both the individual's non-preference attributes and her preferences (Adler, 2016b, 2019, 2022; Adler & Decancq, 2022).

From the policy-P risk profile, we can derive a lottery for individual i over possible lifespans. Let $\mu_{P,i}^l$ be the probability with policy P that i has a "lifespan" of l, that is, lives exactly l periods. This probability can be calculated from the risk profile. The policy-P attribute profile then specifies i's lifetime bundle if she lives exactly l periods, namely: a lifetime bundle consisting in period bundle $b_{P,i}^t$ for periods 1 through l, and then Dead for periods $l+1$ through T.

In short, from the array of individual risk profiles and attribute profiles for each policy P, we can calculate the array of individual lotteries over lifetime bundles for each policy. From this information, we can calculate SU, EPP, and EAP scores for each policy. Yet more simply (because SU, EPP, and EAP satisfy Policy Separability) we can calculate "truncated" SU, EPP, and EAP scores – summing only over affected individuals – and rank policies according to these truncated scores. An individual is "unaffected" if the choice among policies in $\mathbf{P}$ does not change her lottery over lifetime bundles. That is, an unaffected individual has the very same risk profile and attribute profile with each policy in $\mathbf{P}$; affected individuals do not satisfy this condition.

This is a flexible apparatus for thinking about the effect of policies on fatality risks and/or health *and* an individual's material resources (income, consumption, and wealth). Fatality risks figure into the risk profile. Any type of governmental policy that affects fatality risks can be conceptualized as changing individuals' risk profiles, relative to the baseline risk profile. Let $\mathbf{p}_{B,i} = \left(p_{B,i}^{A_i+1}, \ldots, p_{B,i}^{T}\right)$ denote individual i's baseline risk profile ("B" indicating baseline), and let policy P be associated with a list of deltas (changes) to survival probability in the current period and/or future periods: $\left(\Delta p_{P,i}^{A_i+1}, \ldots, \Delta p_{P,i}^{T}\right)$. Perturbing the baseline risk profile by the policy-P deltas, we arrive at the policy-P risk profile. If health quality is one of the attributes in period bundles, policy effects on morbidity as opposed to mortality will

---

and-attribute-profile apparatus is that we can identify how policies change survival probabilities and attributes in each of the possible T periods of i's life, and which period is the present period, not whether we refer to the present period as A_i or (A_i+1).

show up in deltas to individuals' attribute profiles.[27] And, because material resources can also be reflected in attribute profiles (by making income, consumption, or wealth one of the attributes in period bundles), the apparatus is well positioned to capture tradeoffs between fatality risks and/or health, on the one hand, and material resources on the other.

Finally, it should be mentioned that the risk-and-attribute-profile apparatus meshes smoothly with the "survival curve," a standard tool of demography. An individual's baseline risk profile takes the information in her current survival curve (unconditional probability of surviving to each age), and repackages it as a list of conditional survival probabilities.

Adler (2019, chap. 5) builds a simulation model of fatality-risk policy using this apparatus.[28] The affected population consists of 25 equal size age-income cohorts: five age groups (individuals currently aged 20, 30, 40, 50, and 60), crossed with five income groups (individuals with low, moderate, middle, high, and top incomes – corresponding to the 10th, 30th, 50th, 70th, and 90th percentiles of the U.S. income distribution). The period length is 1 year. Baseline risk profiles for each cohort are based upon age-specific death rates taken from a recent U.S. life table, as adjusted to account for the effect of income on survival probability. Attribute profiles describe individual income. Income profiles are constant.[29] Consumption is assumed to be "myopic": individuals consume income when received. Lifetime well-being is the sum of the logarithm of period income.

Adler (2019, chap. 5) considers hypothetical policies, consisting in an average 1-in-100,000 reduction in current fatality risk for individuals in the 25 age-income cohorts: either spread uniformly across the cohorts, or concentrated on the youngest or the poorest individuals. These policies are evaluated using SU; EPP with an "Atkinson" prioritarian SWF[30] with a moderate degree of priority for the worse off ($\gamma = 2$);[31] and BCA.[32] For a given such population-wide risk reduction, *breakeven* average reductions in individuals' current income are calculated. SU breakevens are such that: A policy *P* that produces the stipulated individual risk reduction and an average reduction in current individual income that is less than/equal to/greater than the breakeven income reduction is seen by SU as better than/ equally good as/worse than the status quo. EPP and BCA breakevens are defined analogously. Two possible patterns of cost incidence are considered: uniform cost incidence (individuals incur the same reduction in current income) and proportional cost incidence

[27] For example, health quality might be captured in a numerical value *h* on a 0–1 scale (as in the QALY approach; see Pinto-Prades *et al.*, 2016), with 1 perfect health and 0 a state equivalent to death. Policy effects on an individual's morbidity will then show up as changes to this *h* value in the individual's period bundle for the current period and/or future periods.

[28] The same simulation model was used in the unpublished Adler (2017).

[29] An individual receives the same income in each period of her life, conditional on surviving to its end, which will be low, moderate, middle, high, or top income, depending on the individual's income group.

[30] The Atkinson prioritarian SWF uses the following transformation function: $g(w) = \frac{1}{1-\gamma}w^{1-\gamma}$, with $\gamma > 0$. See Adler (2019, 2022) for further discussion.

[31] Adler (2019, 179, Table 5.3) explains the significance of different values of γ in terms of leaky income transfers. The specific breakeven results in Table 5 below depend upon γ. EPP's upweighting of risk reduction for the young as compared to utilitarianism – as visible in Tables 6 and 7 and discussed below with respect to "Ratio Priority for the Young" – is true regardless of γ. The value of γ determines whether EPP fully neutralizes or reverses the utilitarian preference to confer risk reduction upon richer individuals; see Adler *et al.* (2021, Figure 1) and Tables 6 and 7.

[32] In this simulation, individual *i*'s monetary equivalent for a reduction Δp in current fatality risk plus a reduction Δy in current income was calculated as $(\Delta p)\text{VSL}_i - \Delta y$. See Adler (2019, 295–296).

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

(every individual's current income is reduced by the same fraction of her baseline income, so that members of richer cohorts incur a larger absolute reduction).

Table 5 displays these breakevens. Reading across each row, SU and EPP breakevens differ significantly from each other, and from the BCA breakeven. Adler (2019, chap. 5) explores the sources of these differences. In a nutshell, the three methodologies differ in the relative social value they assign to risk reduction in the 25 cohorts; in the relative social value they assign to income change in the 25 cohorts; and in the social value of risk reduction (SVRR) as compared to income change.

The concept of the SVRR makes precise how SU, EPP, and EAP value risk reduction.[33] SU, EPP, and EAP each assign a score to the baseline array of individual risk and attribute profiles; and a score to the array of individual risk and attribute profiles associated with a given policy *P*. The SVRR concept captures how these scores change as survival probabilities do. SVRR_i^{SU} is the partial derivative of $S^{SU}(\cdot)$ with respect to *i*'s current survival probability, with this partial derivative evaluated at the baseline array of risk and attribute profiles. SVRR_i^{EPP} and SVRR_i^{EAP} are defined analogously: they are the partial derivatives of $S^{EPP}(\cdot)$ and $S^{EAP}(\cdot)$ with respect to *i*'s current survival probability, with these partial derivatives evaluated at the baseline array of risk and attribute profiles. In other words, the SVRRs measure the social value per unit of current risk reduction for individual *i*, for a marginal such reduction – social value according to simple utilitarianism $\left(\text{SVRR}_i^{SU}\right)$, ex post prioritarianism $\left(\text{SVRR}_i^{EPP}\right)$, or ex ante prioritarianism $\left(\text{SVRR}_i^{EAP}\right)$.

Analogous quantities can be defined for changes to future survival probability. In what follows, however, I focus on the SVRR understood in term of current survival probability.

The SVRR_i captures that *portion* of a policy's impact on social value that results from the delta to individual *i*'s current survival probability. Moreover, by comparing SVRR_i to SVRR_j for two individuals *i* and *j*, we can determine the relative social value of risk reduction for the two. Consider a change Δp to someone's current survival probability. That risk change, if accruing to individual *i*, results in a change of social value by approximately $\text{SVRR}_i \times \Delta p$. The very same risk change, accruing instead to individual *j*, results in a change of social value by approximately $\text{SVRR}_j \times \Delta p$. Thus (for a small Δp), the first change in social value is larger than/smaller than/equal to the second iff SVRR_i is larger than/smaller than/equal to SVRR_j.

General formulas for SVRR_i^{SU}, SVRR_i^{EPP}, and SVRR_i^{EAP} are provided in Appendix C. These formulas hold good for any types of attributes included in the attribute profile and any measure of lifetime well-being. The formulas show that SVRR_i^{SU} is the difference between *i*'s baseline expected lifetime well-being, conditional on surviving the period, and her baseline realized well-being if she dies now. SVRR_i^{EPP} is the difference between *i*'s baseline expected *transformed* lifetime well-being, conditional on surviving the period, and her baseline *transformed* realized lifetime well-being if she dies now. SVRR_i^{EAP} is the simple-utilitarian SVRR multiplied by a factor equaling the slope of the transformation function at baseline expected lifetime well-being; this reflects that EAP applies the transformation function to individuals' well-being expectations.

One virtue of the SVRR concept is that it helps to clarify the difference between the SWF methodology and BCA with respect to valuation of fatality-risk reduction. The construct that

[33] On the SVRR concept in the risk-and-attribute-profile apparatus, see Adler *et al.* (2021) and Hammitt and Treich (2022). On the SVRR concept in other setups, see Adler *et al.* (2014) and Hammitt and Treich (2022).

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

Table 5. *Breakeven average individual costs for a policy with an average individual risk reduction of 1-in-100,000*

	SU	EPP	BCA
Uniform risk reduction			
with uniform cost incidence	\$48	\$78	\$91
with proportional cost incidence	\$77	\$159	\$91
Risk reduction for the youngest			
with uniform cost incidence	\$71	\$178	\$132
with proportional cost incidence	\$108	\$360	\$132
Risk reduction for the poorest			
with uniform cost incidence	\$32	\$98	\$15
with proportional cost incidence	\$51	\$201	\$15

BCA employs to value changes in fatality risk is the "value of statistical life" (VSL).[34] VSL_i is the marginal rate of substitution between *i*'s material resources (income, wealth, or consumption) in some period, and *i*'s survival probability in that period. More intuitively: VSL_i is a conversion factor that converts a small change Δp in *i*'s survival probability into an equivalent change in material resources.

VSL_i, understood in terms of current-period survival probability, is the BCA analogue to $SVRR_i$ (Adler *et al.,* 2021). BCA is the sum, across individuals, of their monetary equivalents for a given policy (compensating or equivalent variations) relative to baseline. VSL_i is the change in *i*'s monetary equivalent, per unit of reduction in current-period fatality risk, for a marginal such reduction. Social value, according to BCA, is the sum of monetary equivalents. Thus, VSL_i is the change in social value *as measured by BCA*, per unit of reduction in *i*'s current-period fatality risk, for a marginal such reduction – just as $SVRR_i^{SU}$, $SVRR_i^{EPP}$, and $SVRR_i^{EAP}$ are the changes in social value *as measured by simple utilitarianism, ex post prioritarianism, and ex ante prioritarianism* (respectively) per unit of reduction in *i*'s current-period fatality risk, for a marginal such reduction.

In the simulation model of Adler (2019, chap. 5), VSL_i is specifically equal to $SVRR_i^{SU}$ divided by *i's* expected marginal utility of current income.

Adler (2019, chap. 5) calculates $SVRR_i^{SU}$, $SVRR_i^{EPP}(\gamma = 2)$, and VSL_i for the 25 cohorts, in each case normalized so that 1 is the value for the 60-year-old, low income cohort. The results are displayed in Tables 6–8.

In Table 6, $SVRR_i^{SU}$ decreases moving down within each column (i.e. within each income group as age increases). In this simulation model, the individual's life expectancy remaining (the difference between expected lifespan, conditional on surviving the period, and realized lifespan if the individual dies now) decreases with age – and thus so does $SVRR_i^{SU}$ as individuals become older, holding constant income. $SVRR_i^{SU}$ increases moving right within each row (i.e. within each age group as income increases). Increasing an individual's expected lifespan by a given quantum of expected years produces a larger change in expected well-being, the higher the income that would be earned during those years.

[34] See Adler (2020b) for citations to the literature on VSL.

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

Table 6. *Simple-utilitarian SVRRs*

	Income: Low	Moderate	Middle	High	Top
Age: 20	2.8	3.5	4.0	4.5	5.7
30	2.3	2.9	3.4	3.8	4.8
40	1.8	2.4	2.7	3.1	3.9
50	1.4	1.8	2.1	2.4	3.1
60	1.0	1.3	1.6	1.8	2.3

Table 7. *Ex-post-prioritarian SVRRs*

	Income: Low	Moderate	Middle	High	Top
Age: 20	9.1	7.6	7.0	6.4	5.4
30	5.0	4.2	3.9	3.6	3.0
40	2.9	2.5	2.3	2.1	1.8
50	1.7	1.5	1.4	1.3	1.1
60	1.0	0.9	0.8	0.8	0.7

In Table 7, SVRR_i^{EPP} decreases even more steeply within each column than SVRR_i^{SU}. SVRR_i^{EPP} decreases with age, holding constant income, for two reasons. First, a given increase in current survival probability produces a larger increase in expected lifespan for a younger individual – and thus a greater increase in expected lifetime well-being, holding constant income. This is what drives the decrease of SVRR_i^{SU} within each column; it also affects SVRR_i^{EPP}, but a second factor does too. Younger individuals have a shorter expected lifespan than older individuals – and thus a lower expected level of lifetime well-being. A given increment to lifetime well-being is assigned greater weight by the prioritarian SWF, the lower the level of lifetime well-being at which the increment occurs.

While SVRR_i^{SU} increases within each row, SVRR_i^{EPP} decreases. This pattern results from two, competing factors. The first was already mentioned: holding constant age, a reduction in current fatality risk produces a larger increase in expected lifetime well-being as income increases. The second, countervailing factor is this: as between two individuals of the same age, the one with lower income can expect a lower level of lifetime well-being. Thus, a given increment in her lifetime well-being is assigned greater weight by the prioritarian SWF than the same increment in the lifetime well-being of the richer individual. If the prioritarian SWF is sufficiently concave (as occurs here with the Atkinson-prioritarian SWF and $\gamma = 2$), this second effect will predominate – as can be seen in the row-wise pattern of values in Table 7.

VSL_i (Table 8) has a similar pattern with respect to age as SVRR_i^{SU}. This is because income is modeled as constant within each income group, so that the marginal utility of income is also constant. More realistically, the time path of income is not constant; it tends to rise, then fall, with age. If so, the pattern of VSL_i with respect to age can deviate significantly from SVRR_i^{SU} (Adler, 2020b; Adler *et al.*, 2021).

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

Table 8. *VSL values*

	Income: Low	Moderate	Middle	High	Top
Age: 20	2.8	6.2	9.9	15.6	44.6
30	2.3	5.2	8.3	13.1	37.8
40	1.8	4.2	6.7	10.7	31.0
50	1.4	3.2	5.2	8.4	24.5
60	1.0	2.3	3.9	6.2	18.4

VSL_i's pattern with respect to income is strikingly different from that of $SVRR_i^{SU}$, let alone $SVRR_i^{EPP}$. $SVRR_i^{EPP}$ decreases with income, while VSL_i increases – and much more sharply than $SVRR_i^{SU}$. For example, in the 60-year-old row, $SVRR_i^{SU}$ increases by a factor of 2.3, while VSL_i increases by a factor of 18.4! In Adler (2019, chap. 5) as mentioned, VSL_i equals $SVRR_i^{SU}$ divided by the expected marginal utility of income. Because this denominator decreases with income, VSL_i increases with income more rapidly than $SVRR_i^{SU}$.[35]

Let us return to Table 5, listing breakeven costs for risk-reduction policies, as calculated by SU, EPP, and BCA. The BCA breakevens differ substantially from those for SU and EPP. This occurs for two reasons. The first is differences in the valuation of risk reduction – that is, the differences between VSL_i, $SVRR_i^{SU}$, and $SVRR_i^{EPP}$. The second is differences in how SU, EPP, and BCA value income reductions to the 25 cohorts. SU places a smaller weight on income change for richer individuals, because of the diminishing marginal utility of income: a given change in income produces a smaller change in well-being for a richer person. EPP places a smaller weight on income change for richer individuals for two reasons: the diminishing marginal utility of income *and* because well-being changes to those who are better off are downweighted. BCA assigns equal weight to income changes, regardless of the income level of those whose income changes. As a result, as shown in Table 5, SU breakevens increase by a factor of roughly 1.6 moving from uniform to proportional cost incidence for each of the three patterns of risk reduction, and EPP breakevens increase by even more (roughly doubling), while BCA breakevens are unchanged.

Adler *et al.* (2021) undertakes a detailed theoretical analysis of $SVRR_i$ and VSL_i using the risk-and-attribute-profile apparatus.[36] As in the empirical simulation in Adler (2019, chap. 5), income is the only attribute (thus attribute profiles are income profiles) and consumption is myopic. A headline result of Adler *et al.* (2021) is the finding that both EPP and EAP underwrite the "fair innings" principle. This principle, defended by various public health scholars and philosophers, stipulates that younger individuals should take priority over older individuals with respect to lifesaving policies as a matter of fairness, and not merely because of differential life expectancy gains. Bognar (2015, 254) uses the following thought experiment to crystallize the fair innings principle.

> [Y]ou have only one drug and there are two patients who need it. The only difference between the two patients is their age…. You have to choose between saving: (C) a

[35] The simulation in Adler *et al.* (2021) finds a similar effect.

[36] Adler *et al.* (2021) also undertakes an empirical simulation – calculating $SVRR_i$ and VSL_i for individuals aged 20–100, in five income groups.

20-year old patient who will live for 10 more years if she gets the drug; or (D) a 70-year old patient who will live for 10 more years if she gets the drug.

Both patients would spend the remaining ten years of their life in good health. So there is no difference in expected benefit. The only difference is how much they have already lived when they receive the benefit.

… [According to] the fairness-based argument for the fair innings view, you should … prefer C to D.

Adler *et al.* (2021) demonstrate the following. Let j and k be two individuals with j older than k, and otherwise similarly situated.[37] Although SVRR_i^{SU} often decreases with age (as in the simulation in Adler, 2019, chap. 5, discussed above), this need not hold true. It is possible that $\text{SVRR}_j^{SU} > \text{SVRR}_k^{SU}$. But regardless of the pattern of SVRR_i^{SU} with age, both SVRR_i^{EPP} and SVRR_i^{EAP} give *more* weight to risk reduction for the young than SVRR_i^{SU}. Adler *et al.* (2021) term this "ratio priority for the young."

Ratio Priority for the Young. Let j and k be two individuals with j older than $k\left(A_j > A_k\right)$ and otherwise similarly situated. Then: (1) $\text{SVRR}_k^{EPP}/\text{SVRR}_j^{EPP} > \text{SVRR}_k^{SU}/\text{SVRR}_j^{SU}$; and (2) $\text{SVRR}_k^{EAP}/\text{SVRR}_j^{EAP} > \text{SVRR}_k^{SU}/\text{SVRR}_j^{SU}$. By contrast, VSL_i does not satisfy ratio priority for the young.

The difference in the utilitarian value of risk reduction $\left(\text{SVRR}_i^{SU}\right)$ between young and old reflects the difference in the gain to expected lifetime well-being from reducing a younger versus older person's risk. Prioritarianism, both in the form of EPP and in the form of EAP, provides *greater* relative weight to risk reduction for the young than utilitarianism – and thereby provides a formal expression of the "fair innings" principle.

A second part of the theoretical exercise undertaken by Adler *et al.* (2021) is to do a comparative statics analysis, looking at how the valuation of risk reduction depends upon an individual's baseline income and survival probability. Consider two individuals of the same age, but differing with respect to income in a single period (past, present, or future); with respect to income in all periods; with respect to survival probability in a single period (present or future); or with respect to survival probability in the present period and all future periods. How do SVRR_i^{SU}, SVRR_i^{EPP}, SVRR_i^{EAP}, and VSL_i compare for the two individuals? This comparative-statics exercise confirms significant differences between the four approaches. The prioritarian SVRRs are "history-dependent" (the valuation of present risk reduction changes with past income), while SVRR_i^{SU} and VSL_i are not. VSL_i and SVRR_i^{SU} have the same comparative statics with respect to income but not survival probability; VSL_i and the prioritarian SVRRs have different comparative statics with respect to both income and survival probability. Finally, although the prioritarian SVRRs converge in underwriting

[37] In Adler *et al.* (2021), survival probabilities for each period are given at birth and do not change with age. In this setup, j and k are two individuals who differ in age and have identical income and risk profiles. As Adler *et al.* (2021) explain, all of the results hold true if survival probabilities for past periods are set to 1. That is the approach taken in the statement of the risk-and-attribute-profile approach in the current article: $p_i = 1$ for $t < A_i + 1$. In this setup, two individuals of different ages, j older than k, are similarly situated if they have the same income profiles, and the same survival probabilities for every $t \geq A_j + 1$.

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

the fair innings principle, their comparative statics with respect to income and survival probability are quite different.

I have focused on the risk-and-attribute-profile apparatus for applying the SWF framework to health policy. This is surely *not* the only modeling apparatus for doing so. One-period models, or a different type of multiperiod model, might be employed instead (Adler *et al.,* 2014; 2023; Cookson *et al.,* 2022; Ferranna *et al.,* 2022, 2023; Hammitt & Treich, 2022). One-period models are surely more tractable than the risk-and-attribute-profile apparatus, and a different multiperiod methodology may also be. But the apparatus offers, I believe, an especially powerful lens for bringing into view policy impacts on the different dimensions of lifetime well-being, across a population of individuals differently situated with respect to those dimensions – expected longevity, health, the material resources that enable flourishing, and others – and weighing those impacts in light of social welfare.

Much research is needed to develop the apparatus. Fruitful avenues for inquiry include: having consumption be endogenous rather than myopic; undertaking simulation models and theoretical analysis with both income and health quality as attributes; allowing for stochastic rather than deterministic attribute profiles;[38] taking account of heterogeneous preferences; translating the apparatus to continuous time; and allowing for uncertainty modules that violate Policy Separability.

Acknowledgements. This paper builds on a 2022 conference held by the Brocher Foundation, "Healthy, Wealthy, and Wise – The Ethics of Health Valuation," organized by Nir Eyal (Rutgers University), Samia Hurst (University of Geneva), Lisa A. Robinson (Harvard University), and Daniel Wikler (Harvard University). I thank the Brocher Foundation and conference participants for their support and helpful comments. This special issue was supported by the Brocher Foundation, with supplemental funding from the Rutgers University Center for Population-Level Bioethics, the University of Geneva Institute for Ethics, History, and the Humanities, and the University of Bergen Centre for Ethics and Priority Setting. More information on the Brocher Foundation is available at https://fondation-brocher.ch/.

Competing interest. I declare no competing interests.

References

Adler, Matthew D. 2012. *Well-Being and Fair Distribution: Beyond Cost–Benefit Analysis*. New York: Oxford University Press.

Adler, Matthew D. 2016a. "Benefit–Cost Analysis and Distributional Weights: An Overview." *Review of Environmental Economics and Policy*, *10*: 264–285.

Adler, Matthew D. 2016b. "Extended Preferences." In Adler, Matthew D., and Marc Fleurbaey (Eds.) *The Oxford Handbook of Well-Being and Public Policy*, pp. 476–517. New York: Oxford University Press.

Adler, Matthew D. 2017. "A Better Calculus for Regulators: From Cost–Benefit Analysis to the Social Welfare Function." Working paper, Duke Law School Public Law and Legal Theory Series No. 2017-19. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2923829.

Adler, Matthew D. 2019. *Measuring Social Welfare: An Introduction*. New York: Oxford University Press.

Adler, Matthew D. 2020a. "Social Welfare Functions." In Norheim, Ole F., Ezekiel J. Emanuel, and Joseph Millum (Eds.) *Global Health Priority-Setting: Beyond Cost-Effectiveness*, pp. 123–141. New York: Oxford University Press.

Adler, Matthew D. 2020b. "What Should We Spend to Save Lives in a Pandemic? A Critique of the Value of Statistical Life." *Covid Economics*, *33*: 1–45.

[38] See Adler (forthcoming, chap. 5). Stochastic attribute profiles would, inter alia, allow the framework to take account of uncertainty about policy effects on morbidity.

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

Adler, Matthew D. 2022. "Theory of Prioritarianism." In Adler, Matthew D., and Ole F. Norheim (Eds.) *Prioritarianism in Practice*, pp. 37–127. Cambridge: Cambridge University Press.

Adler, Matthew D. Forthcoming. *Risk, Death, and Well-Being: The Ethical Foundations of Fatality Risk Regulation*. New York: Oxford University Press.

Adler, Matthew D., Richard Bradley, Maddalena Ferranna, Marc Fleurbaey, James Hammitt, Rémi Turquier, and Alex Voorhoeve. 2023. "How to Balance Lives and Livelihoods in a Pandemic." In Savulescu, Julian, and Dominic Wilkinson (Eds.) *Pandemic Ethics: From COVID-19 to Disease X*, pp. 189–209. Oxford: Oxford University Press.

Adler, Matthew D., and Koen Decancq. 2022. "Well-Being Measurement." In Adler, Matthew D., and Ole F. Norheim (Eds.) *Prioritarianism in Practice*, pp. 128–171. Cambridge: Cambridge University Press.

Adler, Matthew D., Maddalena Ferranna, James K. Hammitt, and Nicolas Treich. 2021. "Fair Innings? The Utilitarian and Prioritarian Value of Risk Reduction over a Whole Lifetime." *Journal of Health Economics*, *75*: 102412.

Adler, Matthew D., James K. Hammitt, and Nicolas Treich. 2014. "The Social Value of Mortality Risk Reduction: VSL versus the Social Welfare Function Approach." *Journal of Health Economics*, *35*: 82–93.

Adler, Matthew D., and Nils Holtug. 2019. "Prioritarianism: A Response to Critics." *Politics, Philosophy & Economics*, *18*: 101–144.

Blackorby, Charles, Walter Bossert, and David Donaldson. 2005. *Population Issues in Social Choice Theory, Welfare Economics, and Ethics*. Cambridge: Cambridge University Press.

Bleichrodt, Han, Enrico Diecidue, and John Quiggin. 2004. "Equity Weights in the Allocation of Health Care: The Rank-Dependent QALY Model." *Journal of Health Economics*, *23*: 157–171.

Boadway, Robin W. 2016. "Cost–Benefit Analysis." In Adler, Matthew D., and Marc Fleurbaey (Eds.) *The Oxford Handbook of Well-Being and Public Policy*, pp. 47–81. New York: Oxford University Press.

Bognar, Greg. 2015. "Fair Innings." *Bioethics*, *29*: 251–261.

Bossert, Walter, and John A. Weymark. 2004. "Utility in Social Choice." In Barberà, Salvador, Peter J. Hammond, and Christian Seidl (Eds.) *Handbook of Utility Theory*, vol. 2 (Extensions), pp. 1099–1177. Boston: Kluwer Academic.

Botzen, W. J. Wouter, and Jeroen C. J. M. van den Bergh. 2014. "Specifications of Social Welfare in Economic Studies of Climate Policy: Overview of Criteria and Related Policy Insights." *Environmental and Resource Economics*, *58*: 1–33.

Cookson, Richard, Ole F. Norheim, and Ieva Skarda. 2022. "Prioritarianism and Health Policy." In Adler, Matthew D., and Ole F. Norheim (Eds.) *Prioritarianism in Practice*, pp. 260–316. Cambridge: Cambridge University Press.

Dolan, Paul. 1998. "The Measurement of Individual Utility and Social Welfare." *Journal of Health Economics*, *17*: 39–52.

Ferranna, Maddalena, James K. Hammitt, and Matthew D. Adler. 2023. "Age and the Value of Life." In Bloom, David E., Alfonso Sousa-Poza, and Uwe Sunde (Eds.) *The Routledge Handbook of the Economics of Ageing*, pp. 566–577. Abingdon: Routledge.

Ferranna, Maddalena, J. P. Sevilla, and David E. Bloom. 2022. "Prioritarianism and the COVID-19 Pandemic." In Adler, Matthew D., and Ole F. Norheim (Eds.) *Prioritarianism in Practice*, pp. 572–650. Cambridge: Cambridge University Press.

Fleurbaey, Marc. 2010. "Assessing Risky Social Situations." *Journal of Political Economy*, *118*: 649–680.

Fleurbaey, Marc. 2016. "Equivalent Income." In Adler, Matthew D., and Marc Fleurbaey (Eds.), *The Oxford Handbook of Well-Being and Public Policy*, pp. 453–475. New York: Oxford University Press.

Fleurbaey, Marc, and Rossi Abi-Rafeh. 2016. "The Use of Distributional Weights in Benefit–Cost Analysis: Insights from Welfare Economics." *Review of Environmental Economics and Policy*, *10*: 286–307.

Hammitt, James K., and Nicolas Treich. 2022. "Prioritarianism and Fatality Risk Regulation." In Adler, Matthew D., and Ole F. Norheim (Eds.) *Prioritarianism in Practice*, pp. 317–359. Cambridge: Cambridge University Press.

Hougaard, Jens Leth, Juan D. Moreno-Ternero, and Lars Peter Østerdal. 2013. "A New Axiomatic Approach to the Evaluation of Population Health." *Journal of Health Economics*, *32*: 515–523.

Mongin, Philippe, and Marcus Pivato. 2016. "Social Evaluation under Risk and Uncertainty." In Adler, Matthew D., and Marc Fleurbaey (Eds.) *The Oxford Handbook of Well-Being and Public Policy*, pp. 711–745. New York: Oxford University Press.

Nurmi, Väinö, and Heini Ahtiainen. 2018. "Distributional Weights in Environmental Valuation and Cost–Benefit Analysis: Theory and Practice." *Ecological Economics*, *150*: 217–228.

Østerdal, Lars Peter. 2005. "Axioms for Health Care Resource Allocation." *Journal of Health Economics*, *24*: 679–702.

Pinto-Prades, Jose-Luís, Carmen Herrero, and Jose María Abellán. 2016. "QALY-Based Cost-Effectiveness Analysis." In Adler, Matthew D., and Marc Fleurbaey (Eds.) *The Oxford Handbook of Well-Being and Public Policy*, pp. 160–192. New York: Oxford University Press.

Tuomala, Matti. 2016. *Optimal Redistributive Taxation*. Oxford: Oxford University Press.

U.S. Office of Management and Budget. 2023. "Circular No. A-4, November 9, 2023." https://www.whitehouse.gov/wp-content/uploads/2023/11/CircularA-4.pdf.

Weymark, John. 2016. "Social Welfare Functions." In Adler, Matthew D., and Marc Fleurbaey (Eds.) *The Oxford Handbook of Well-Being and Public Policy*, pp. 126–159. New York: Oxford University Press.

Williams, Alan. 1997. "Intergenerational Equity: An Exploration of the 'Fair Innings' Argument." *Health Economics*, *6*: 117–132.

Appendix

A. Uncertainty axioms

Here are formal statements of the uncertainty axioms stated informally in the text.

> *Ex Ante Pareto*. (1) *Ex Ante Pareto Indifference*. If $\sum_x \pi_P(x)\mathbf{w}_i(x) = \sum_x \pi_{P^*}(x)\mathbf{w}_i(x)$ for all i, then $P \sim^{\mathbf{P}} P^*$. (2) *Ex Ante Strong Pareto*. If $\sum_x \pi_P(x)\mathbf{w}_i(x) \geq \sum_x \pi_{P^*}(x)\mathbf{w}_i(x)$ for all i, and $\sum_x \pi_P(x)\mathbf{w}_j(x) > \sum_x \pi_{P^*}(x)\mathbf{w}_j(x)$ for at least one j, $P \succ^{\mathbf{P}} P^*$.

> *Dominance*. If: for every x, x^* such that $\pi_P(x) > 0$ and $\pi_{P^*}(x^*) > 0$, $\mathbf{w}(x) \succ \mathbf{w}(x^*)$ according to some SWF, then: the uncertainty module for that SWF should be such that $P \succ^{\mathbf{P}} P^*$.

The statement of Policy Separability will employ the following notation. $L_{P,i}$ denotes the lottery over well-being levels for individual i that results from policy P. With v a real number, $L_{P,i}(v) = \sum_{x:\mathbf{w}_i(x)=v} \pi_P(x)$, that is, $L_{P,i}(v)$ is the probability to individual i of well-being level v with policy P. $L_{P,i} = L_{P^*,i}$ indicates that i faces the same well-being lottery with policies P and P^*, that is: for every real number v, $L_{P,i}(v) = L_{P^*,i}(v)$.

> *Policy Separability*. (1) If $L_{P,i} = L_{P^*,i}$ for all i, then $P \sim^{\mathbf{P}} P^*$. (2) Let $\mathbf{M}$ be any subset of $\mathbf{I}$, and let $\mathbf{M}^+ = \mathbf{I} \setminus \mathbf{M}$ (all individuals not in $\mathbf{M}$). Assume P, P^*, P^+, P^{++} are as follows. For all $i \in \mathbf{M}$, $L_{P,i} = L_{P^*,i}$ and $L_{P+,i} = L_{P++,i}$. For all $j \in \mathbf{M}^+$, $L_{P,j} = L_{P+,j}$ and $L_{P^*,j} = L_{P++,j}$. Then $P \succsim^{\mathbf{P}} P^*$ iff $P^+ \succsim^{\mathbf{P}} P^{++}$.

As mentioned in the text, the second prong of Policy Separability implies the first (but not vice versa). To see the implication, let P^+ and P^{++} be "degenerate" policies each of which yields the same well-being level v_i' for sure for each individual i: $L_{P+,i}(v_i') = L_{P++,i}(v_i') = 1$. Assume that policies P and P^* meet the antecedent condition for the first prong, namely, $L_{P,i} = L_{P^*,i}$ for all i. Assume that the second prong holds true: $P \succsim^{\mathbf{P}} P^*$ iff $P^+ \succsim P^{++}$. Because the ranking of degenerate policies is consistent with the SWF and by Pareto Indifference for well-being vectors, $P^+ \sim^{\mathbf{P}} P^{++}$. Thus, $P \sim^{\mathbf{P}} P^*$.[39]

[39] The policies P, P^*, P^+, and P^{++} referred to by the second prong of Policy Separability could be any four policies meeting the conditions stated there. Thus, it is true not only that $P \succsim^{\mathbf{P}} P^*$ iff $P^+ \succsim^{\mathbf{P}} P^{++}$ but also that $P^* \succsim^{\mathbf{P}} P$ iff $P^{++} \succsim^{\mathbf{P}} P^+$. Putting these two biconditionals together, we have that $P \sim^{\mathbf{P}} P^*$ iff $P^+ \sim^{\mathbf{P}} P^{++}$.

B. Bundle lotteries

As stated in the text, the scores assigned to policies by SU, EPP, and EAP (formulas given in Section 2) can be restated as a function of the array of individual bundle lotteries (formulas given at the beginning of Section 3). To see this, recall that $\mathbf{w}_i(x) = w(b_i(x))$, $b_i(x)$ the bundle of individual i in outcome x. Let $\rho_{P,i}(b)$ denote the probability that individual i receives bundle b with policy P. $\rho_{P,i}(b) = \sum_{x:b_i(x)=b} \pi_P(x)$.

Then $S^{SU}(P) = \sum_{i=1}^{N}\sum_x \pi_P(x)\mathbf{w}_i(x) = \sum_{i=1}^{N}\sum_b \rho_{P,i}(b)w(b)$. $S^{EPP}(P) = \sum_{i=1}^{N}\sum_x \pi_P(x)g(\mathbf{w}_i(x)) = \sum_{i=1}^{N}\sum_b \rho_{P,i}(b)g(w(b))$. $S^{EAP}(P) = \sum_{i=1}^{N} g\left(\sum_x \pi_P(x)\mathbf{w}_i(x)\right) = \sum_{i=1}^{N} g\left(\sum_b \rho_{P,i}(b)w(b)\right)$.

It was noted in Section 3 that policies can be ranked according to "truncated" SU, EPP, and EAP scores – summing only over affected individuals. To see this, let **A(P)**, the "affected individuals," be a subset of **I** such that: $i \notin \mathbf{A}(\mathbf{P})$ iff for all P, P^* in **P**, $\rho_{P,i}(b) = \rho_{P^*,i}(b)$ for every b. Thus, if i is not a member of $\mathbf{A}(\mathbf{P})$, the following are true for any P, P^*: (1) $\sum_b \rho_{P,i}(b)w(b) = \sum_b \rho_{P^*,i}(b)w(b)$; (2) $\sum_b \rho_{P,i}(b)g(w(b)) = \sum_b \rho_{P^*,i}(b)g(w(b))$; and (3) $g\left(\sum_b \rho_{P,i}(b)w(b)\right) = g\left(\sum_b \rho_{P^*,i}(b)w(b)\right)$. Therefore, (1) $S^{SU}(P) \geq S^{SU}(P^*)$ iff $\sum_{i\in\mathbf{A}(\mathbf{P})}\sum_b \rho_{P,i}(b)w(b) \geq \sum_{i\in\mathbf{A}(\mathbf{P})}\sum_b \rho_{P^*,i}(b)w(b)$; (2) $S^{EPP}(P) \geq S^{EPP}(P^*)$ iff $\sum_{i\in\mathbf{A}(\mathbf{P})}\sum_b \rho_{P,i}(b)g(w(b)) \geq \sum_{i\in\mathbf{A}(\mathbf{P})}\sum_b \rho_{P^*,i}(b)g(w(b))$; and (3) $S^{EAP}(P) \geq S^{EAP}(P^*)$ iff $\sum_{i\in\mathbf{A}(\mathbf{P})} g\left(\sum_b \rho_{P,i}(b)w(b)\right) \geq \sum_{i\in\mathbf{A}(\mathbf{P})} g\left(\sum_b \rho_{P^*,i}(b)w(b)\right)$

C. The risk and attribute profile apparatus

Individual i has a current age A_i; that is, the number of the current period in i's life is $A_i + 1$. T is the maximum possible length of life (number of periods). A policy P endows individual i with a risk profile $\mathbf{p}_{P,i} = (p_{P,i}^{A_i+1}, \ldots, p_{P,i}^{T})$, with $p_{P,i}^t$ the probability with policy P that i survives to the end of period t, conditional on being alive at the beginning of that period. P also endows i with an attribute profile $\mathbf{b}_{P,i} = (b_{P,i}^1, \ldots, b_{P,i}^T)$, with $b_{P,i}^t$ the bundle of attributes that i receives in period t, conditional on surviving to the end of period t.

The individual's policy-specific risk and attribute profiles, in turn, determine her lottery over lifetime bundles. Let $\mu_{P,i}^l$ denote the probability that individual i lives exactly l periods. If $l < A_i$, $\mu_{P,i}^l = 0$. If $l = A_i$, $\mu_{P,i}^l = 1 - p_{P,i}^{A_i+1}$. Finally, if $l > A_i$, $\mu_{P,i}^l = \left(\prod_{t=A_i+1}^{l} p_{P,i}^t\right)(1 - p_{P,i}^{l+1})$. For a given longevity l, her policy-P sequence of period bundles is just $b_{P,i}^1, \ldots, b_{P,i}^l$ and then Dead for periods $l+1$ to T.

Let $W_{P,i}^l$ denote individual i's lifetime well-being if she lives exactly l periods with her policy-P attribute profile. $W_{P,i}^l = w(b_{P,i}^1, \ldots, b_{P,i}^l, Dead, \ldots, Dead)$. Then $S^{SU}(P) = \sum_{i=1}^{N}\sum_{t=A_i}^{T} \mu_{P,i}^t W_{P,i}^t$. $S^{EPP}(P) = \sum_{i=1}^{N}\sum_{t=A_i}^{T} \mu_{P,i}^t g(W_{P,i}^t)$. $S^{EAP}(P) = \sum_{i=1}^{N} g\left(\sum_{t=A_i}^{T} \mu_{P,i}^t W_{P,i}^t\right)$. This shows that $S(P)$ can be written as a function of the population-wide array of risk and attribute profiles with P.

We can now define the SVRR. Let **H** denote the population-wide array of risk and attribute profiles in the baseline. With $S(\cdot) = S^{SU}(\cdot)$, $S^{EPP}(\cdot)$, or $S^{EAP}(\cdot)$, SVRR_i is the partial derivative of $S(\cdot)$ with respect to i's current survival probability, this partial derivative

https://doi.org/10.1017/bca.2024.6 Published online by Cambridge University Press

evaluated at $\mathbf{H}$. That is, $SVRR_i = \left(\frac{\partial S}{\partial p_i^{A_i+1}}\right)(\mathbf{H})$. Because $S^{SU}(\cdot)$, $S^{EPP}(\cdot)$, and $S^{EAP}(\cdot)$ are each additive across individuals, $\frac{\partial S}{\partial p_i^{A_i+1}}(\mathbf{H})$ is a function just of individual i's baseline risk and attribute profiles. Thus, SVRR_i can be written as follows: $SVRR_i = \frac{\partial S}{\partial p_i^{A_i+1}}(\mathbf{p}_{B,i}, \mathbf{b}_{B,i})$.

The text notes that SVRR_i captures that portion of a policy's impact on social value that results from the delta to individual i's current survival probability. To see why, assume (for simplicity) that each period bundle consists in a single numerically measurable attribute. (What follows generalizes to the case of period bundles with multiple numerically measurable attributes.) Assume that a policy changes individual i's current survival probability by Δp_i relative to baseline, as well as perhaps changing her future survival probability and her current and future attributes. Then, by the total differential approximation from calculus, the resultant change in social welfare is approximately the sum across individuals of $\mathrm{SVRR}_i \times \Delta p_i$ plus analogous terms for the changes to future survival probability and to individuals' current and future attributes. Let Δp_i^t be the change to individual i's future survival probability in period t, $t > A_i + 1$; and let Δb_i^s be the change to individual i's current or future attributes in period s, $s \geq A_i + 1$. Then, with ΔS the resultant change to social welfare,

$$\Delta S \approx \sum_{i=1}^{N}\left(SVRR_i \times \Delta p_i + \sum_{t=A_i+2}^{T}\frac{\partial S}{\partial p_i^t}(\mathbf{p}_{B,i}, \mathbf{b}_{B,i}) \times \Delta p_i^t + \sum_{s=A_i+1}^{T}\frac{\partial S}{\partial b_i^s}(\mathbf{p}_{B,i}, \mathbf{b}_{B,i}) \times \Delta b_i^s\right).$$

It is straightforward to show that $SVRR_i^{SU} = -W_{B,i}^{A_i} + \sum_{t=A_i+1}^{T}\frac{\mu_{B,i}^t}{p_{B,i}^{A_i+1}}W_{B,i}^t$. To see this, note that $\frac{\partial \mu_i^t}{\partial p_i^{A_i+1}}(\mathbf{p}_{B,i}, \mathbf{b}_{B,i}) = -1$ for $t = A_i$ and $\frac{\mu_{B,i}^t}{p_{B,i}^{A_i+1}}$ for $t > A_i$. Similarly, $SVRR_i^{EPP} = -g\left(W_{B,i}^{A_i}\right) + \sum_{t=A_i+1}^{T}\frac{\mu_{B,i}^t}{p_{B,i}^{A_i+1}}g\left(W_{B,i}^t\right)$. Finally, $SVRR_i^{EAP} = g'\left(\sum_{t=A_i}^{T}\mu_{B,i}^t W_{B,i}^t\right) \times SVRR_i^{SU}$.

Cite this article: Adler, Matthew D. 2024. "Social Welfare Functions and Health Policy: A New Approach." *Journal of Benefit-Cost Analysis* 15: 26–51, doi:10.1017/bca.2024.6